

LISTING OF THE CLAIMS

Claims 1-50 were originally pending. Please amend claims 27 and 41. No claims are canceled or withdrawn. Accordingly, claims 1-50 remain pending.

The following listing of claims replaces all prior versions and listings of claims in the application.

1. (Previously presented) A computer-implemented method for related term suggestion, the method comprising:

generating term clusters as a function of calculated similarity of term vectors, each term vector being generated from search results associated with a set of high frequency of occurrence (FOO) historical queries previously submitted to a search engine; and

responsive to receiving a term/phrase from an entity, evaluating the term/phrase in view of terms/phrases in the term clusters to identify one or more related term suggestions.

2. (Previously presented) The method of claim 1, and wherein a multi-sense query comprises the term/phrase.

3. (Previously presented) The method of claim 1, and wherein the entity is a computer-program application and/or an end-user.

4. (Previously presented) The method of claim 1, further comprising determining the calculated similarity as follows:

$$\text{sim}(q_j, q_k) = \sum_{i=1}^d w_{ij} \cdot w_{ik} ;$$

wherein d represents vector dimension, q represents a query, k is a dimension index, and wherein weight w for the i^{th} vector's j^{th} term is calculated as follows:

$$w_{ij} = TF_{ij} \times \log(N / DF_j) ; \text{ and}$$

wherein TF_{ij} represents term frequency, N is a total number of query terms, and DF_j is a number of extracted feature records that contain term j .

5. (Previously presented) The method of claim 1, further comprising:
collecting historic query terms from a query log; and
determining ones of the historic query terms with a high FOO.
6. (Previously presented) The method of claim 1, further comprising before creating the term clusters:
reducing dimensionality of the term vectors; and
normalizing the term vectors.

7. (Previously presented) The method of claim 1, wherein evaluating further comprises:

identifying a match between the term/phrase and term(s)/phrase(s) from one or more term clusters; and

responsive to identifying, generating related term suggestion(s) comprising the term(s)/phrase(s).

8. (Previously presented) The method of claim 7, wherein the related term suggestion(s) further comprise for each term/phrase of the term(s)/phrase(s), a frequency of occurrence value indicating a number of times the term/phrase occurs in a set of mined historical queries.

9. (Previously presented) A method as recited in claim 1, wherein generating the term clusters further comprises:

sending respective ones of the high FOO historical queries to the search engine to obtain the search results;

extracting features from at least a subset of search results corresponding to the respective ones; and

producing term vectors from the features as a function of term and inverted document frequencies.

10. (Previously presented) The method of claim 9, and wherein the features comprise a title, description, and/or context for the respective ones of the high FOO historical query terms.

11. (Previously presented) The method of claim 9, and wherein the respective ones comprise top ranked ones of the search results.

12. (Previously presented) The method of claim 1, wherein the term clusters are a first set of term clusters, and wherein the method further comprises:

determining that there is no match between the term/phrase and the terms/phrases; and

responsive to the determining:

making a second set of term clusters from calculated similarity of term vectors, each term vector being generated from search results associated with a set of low FOO historical queries previously submitted to the search engine; and

evaluating the term/phrase in view of terms/phrases of the second set of term clusters to identify one or more related term suggestions.

13. (Previously presented) The method of claim 12, wherein making further comprises:

identifying the low FOO historical queries from historical queries mined from a query log;

sending respective ones of at least a subset of the low FOO historical queries to the search engine to obtain search results;

extracting features from at least a subset of search results; and

producing the term vectors from the features as a function of term and inverted document frequencies.

14. (Previously presented) The method of claim 13, and further comprising after clustering:

determining that there is no match between the term/phrase and term(s)/phrase(s) from the first set of term clusters, the first set being based on high FOO historical queries; and

responsive to the determining, identifying a match between the term/phrase and term(s)/phrase(s) from one or more of the second set of term clusters, the second set being based on low FOO historical queries; and

responsive to identifying, generating related term suggestion(s) comprising the term(s)/phrase(s).

15. (Previously presented) A tangible computer-readable data storage medium comprising computer-executable instructions for:

generating term clusters as a function of calculated similarity of term vectors, each term vector being generated from search results associated with a set of high frequency of occurrence (FOO) historical queries previously submitted to a search engine; and

responsive to receiving a term/phrase from an entity, evaluating the term/phrase in view of terms/phrases in the term clusters to identify one or more related term suggestions.

16. (Previously presented) The computer-readable data storage medium of claim 15, and wherein a multi-sense query comprises the term/phrase.

17. (Previously presented) The computer-readable data storage medium of claim 15, and wherein the entity is a computer-program application and/or an end-user.

18. (Previously presented) The computer-readable data storage medium of claim 15, further comprising computer-executable instructions for determining the calculated similarity as follows:

$$\text{sim}(q_j, q_k) = \sum_{i=1}^d w_{ij} \cdot w_{ik} ;$$

wherein d represents vector dimension, q represents a query, k is a dimension index, and wherein weight w for the i^{th} vector's j^{th} term is calculated as follows:

$$w_{ij} = TF_{ij} \cdot \log(N / DF_j) ; \text{ and}$$

wherein TF_{ij} represents term frequency, N is a total number of query terms, and DF_j is a number of extracted feature records that contain term j .

19. (Previously presented) The computer-readable data storage medium of claim 15, further comprising computer-executable instructions for:

collecting historic query terms from a query log; and

determining ones of the historic query terms with a high FOO.

20. (Previously presented) The computer-readable data storage medium of claim 15, before creating the term clusters, further comprising computer-executable instructions for:

- reducing dimensionality of the term vectors; and
- normalizing the term vectors.

21. (Previously presented) The computer-readable data storage medium of claim 15, wherein evaluating further comprises computer-executable instructions for:

- identifying a match between the term/phrase and term(s)/phrase(s) from one or more term clusters; and

- responsive to identifying, generating related term suggestion(s) comprising the term(s)/phrase(s).

22. (Previously presented) The computer-readable data storage medium of claim 21, wherein the related term suggestion(s) further comprise for each term/phrase of the term(s)/phrase(s), a frequency of occurrence value indicating a number of times the term/phrase occurs in a set of mined historical queries.

23. (Previously presented) The computer-readable data storage medium of claim 15, wherein generating the term clusters further comprises computer-executable instructions for:

 sending respective ones of the high FOO historical queries to the search engine to obtain the search results;

 extracting features from at least a subset of search results corresponding to the respective ones; and

 producing term vectors from the features as a function of term and inverted document frequencies.

24. (Previously presented) The computer-readable data storage medium of claim 23, and wherein the features comprise a title, description, and/or context for the respective ones of the high FOO historical query terms.

25. (Previously presented) The computer-readable data storage medium of claim 23, and wherein the respective ones comprise top ranked ones of the search results.

26. (Previously presented) The computer-readable data storage medium of claim 15, wherein the term clusters are a first set of term clusters, and wherein the computer-executable instructions further comprise instructions for:

determining that there is no match between the term/phrase and the terms/phrases; and

responsive to the determining:

making a second set of term clusters from calculated similarity of term vectors, each term vector being generated from search results associated with a set of low FOO historical queries previously submitted to the search engine; and

evaluating the term/phrase in view of terms/phrases of the second set of term clusters to identify one or more related term suggestions.

27. (Currently amended) The computer-readable data storage medium of claim 26, wherein making further comprises computer-executable instructions for:

identifying the low FOO historical queries from historical queries mined from a query log;

sending respective ones of at least a subset of the low FOO historical queries to the search engine to obtain search results;

extracting features from at least a subset of search results; and

producing the term vectors from the features as a function of term and inverted ~~term~~ document frequencies.

28. (Previously presented) The computer-readable data storage medium of claim 27, and further comprising computer-executable instructions, after clustering, for:

determining that there is no match between the term/phrase and term(s)/phrase(s) from the first set of term clusters, the first set being based on high FOO historical queries; and

responsive to the determining, identifying a match between the term/phrase and term(s)/phrase(s) from one or more of the second set of term clusters, the second set being based on low FOO historical queries; and

responsive to identifying, generating related term suggestion(s) comprising the term(s)/phrase(s).

29. (Original) A computing device comprising:

a processor; and

a memory coupled to the processor, the memory comprising computer-program instructions executable by the processor for:

generating term clusters as a function of calculated similarity of term vectors, each term vector being generated from search results associated with a set of high frequency of occurrence (FOO) historical queries previously submitted to a search engine; and

responsive to receiving a term/phrase from an entity, evaluating the term/phrase in view of terms/phrases in the term clusters to identify one or more related term suggestions.

30. (Previously presented) The computing device of claim 29, and wherein a multi-sense query comprises the term/phrase.

31. (Previously presented) The computing device of claim 29, and wherein the entity is a computer-program application and/or an end-user.

32. (Previously presented) The computing device of claim 29, further comprising computer-executable instructions for determining the calculated similarity as follows:

$$\text{sim}(q_j, q_k) = \sum_{i=1}^d w_{ij} \cdot w_{ik} ;$$

wherein d represents vector dimension, q represents a query, k is a dimension index, and wherein weight w for the i^{th} vector's j^{th} term is calculated as follows:

$$w_{ij} = TF_{ij} \times \log(N / DF_j) ; \text{ and}$$

wherein TF_{ij} represents term frequency, N is a total number of query terms, and DF_j is a number of extracted feature records that contain term j .

33. (Previously presented) The computing device of claim 29, further comprising computer-executable instructions for:

collecting historic query terms from a query log; and

determining ones of the historic query terms with a high FOO.

34. (Previously presented) The computing device of claim 29, before creating the term clusters, further comprising computer-executable instructions for:

reducing dimensionality of the term vectors; and
normalizing the term vectors.

35. (Previously presented) The computing device of claim 29, wherein evaluating further comprises computer-executable instructions for:

identifying a match between the term/phrase and term(s)/phrase(s) from one or more term clusters; and

responsive to identifying, generating related term suggestion(s) comprising the term(s)/phrase(s).

36. (Previously presented) The computing device of claim 35, wherein the related term suggestion(s) further comprise for each term/phrase of the term(s)/phrase(s), a frequency of occurrence value indicating a number of times the term/phrase occurs in a set of mined historical queries.

37. (Previously presented) The computing device of claim 29, wherein generating the term clusters further comprises computer-executable instructions for:

sending respective ones of the high FOO historical queries to the search engine to obtain the search results;

extracting features from at least a subset of search results corresponding to the respective ones; and

producing term vectors from the features as a function of term and inverted document frequencies.

38. (Previously presented) The computing device of claim 37, and wherein the features comprise a title, description, and/or context for the respective ones of the high FOO historical query terms.

39. (Previously presented) The computing device of claim 37, and wherein the respective ones comprise top ranked ones of the search results.

40. (Previously presented) The computing device of claim 29, wherein the term clusters are a first set of term clusters, and wherein the computer-executable instructions further comprise instructions for:

determining that there is no match between the term/phrase and the terms/phrases; and

responsive to the determining:

making a second set of term clusters from calculated similarity of term vectors, each term vector being generated from search results associated with a set of low FOO historical queries previously submitted to the search engine; and

evaluating the term/phrase in view of terms/phrases of the second set of term clusters to identify one or more related term suggestions.

41. (Currently amended) The computing device of claim 40, wherein making further comprises computer-executable instructions for:

identifying the low FOO historical queries from historical queries mined from a query log;

sending respective ones of at least a subset of the low FOO historical queries to the search engine to obtain search results;

extracting features from at least a subset of search results; and

producing the term vectors from the features as a function of term and inverted ~~term~~ document frequencies.

42. (Previously presented) The computing device of claim 41, and further comprising computer-executable instructions, after clustering, for:

determining that there is no match between the term/phrase and term(s)/phrase(s) from the first set of term clusters, the first set being based on high FOO historical queries; and

responsive to the determining, identifying a match between the term/phrase and term(s)/phrase(s) from one or more of the second set of term clusters, the second set being based on low FOO historical queries; and

responsive to identifying, generating related term suggestion(s) comprising the term(s)/phrase(s).

43. (Original) A computing device comprising:

generating means to generate term clusters as a function of calculated similarity of term vectors, each term vector being generated from search results associated with a set of high frequency of occurrence (FOO) historical queries previously submitted to a search engine; and

responsive to receiving a term/phrase from an entity, evaluating means to evaluate the term/phrase in view of terms/phrases in the term clusters to identify one or more related term suggestions.

44. (Previously presented) The computing device of claim 43, and wherein a multi-sense query comprises the term/phrase.

45. (Previously presented) The computing device of claim 43, and wherein the entity is a computer-program application and/or an end-user.

46. (Previously presented) The computing device of claim 43, and further comprising:

collecting means to collect historic query terms from a query log; and

determining means to determine ones of the historic query terms with a high FOO.

47. (Previously presented) The computing device of claim 43, wherein the evaluating means further comprise:

identifying means to identify a match between the term/phrase and term(s)/phrase(s) from one or more term clusters; and

responsive to identifying, generating means to generate related term suggestion(s) comprising the term(s)/phrase(s).

48. (Previously presented) The computing device of claim 43, wherein the generating means to generate the term clusters further comprise:

sending means to send respective ones of the high FOO historical queries to the search engine to obtain the search results;

extracting means to extract features from at least a subset of search results corresponding to the respective ones; and

producing means to produce term vectors from the features.

49. (Original) A computing device as recited in claim 43, wherein the term clusters are a first set of term clusters, and wherein the computing device further comprises:

determining means to determine that there is no match between the term/phrase and the terms/phrases; and

responsive to the determining:

making means to make a second set of term clusters from calculated similarity of term vectors, each term vector being generated from search results associated with a set of low FOO historical queries previously submitted to the search engine; and

evaluating means to evaluate the term/phrase in view of terms/phrases of the second set of term clusters to identify one or more related term suggestions.

50. (Previously presented) The computing device of claim 49, and further comprising:

calculating means to calculate that there is no match between the term/phrase and term(s)/phrase(s) from the first set of term clusters, the first set being based on high FOO historical queries; and

responsive to the calculating, identifying means to identify a match between the term/phrase and term(s)/phrase(s) from one or more of the second set of term clusters, the second set being based on low FOO historical queries; and

responsive to identifying, generating means to generate related term suggestion(s) comprising the term(s)/phrase(s).